W271_Lab3_Final_Project

Shih Yu Chang

December 10, 2016

In this report, I try to apply a new package in R, named as Multivariate Time Series (MTS), developed by Ruey S Tsay at The University of Chicago Booth School of Business. MTS is a general package for analyzing multivariate linear time series and estimating multivariate volatility models. There are two parts of this report. The first part is to study the effect of 1998 financial aid act, *Higher Education Amendments of 1998*, in university/college enrollment rate with respect to sex. The second portion is to predict university/college enrollment rate.

Part I: Effect of Financial Aid with respect to Sex for University/College Enrollment

Higher Education Amendments of 1998 renamed SSIG as LEAP, created GEAR UP, suspended student aid eligibility for drug convictions, and added Extended Repayment. (P.L. 105-244, October 7, 1998). The changes including following: (1) Cut Stafford loan interest rates by 0.80%. Consolidation loans round up to the nearest 1/8th of a percent (previously whole percent) and capped at 8.25%. (2) Cost of attendance may now include the cost of a personal computer. (3) Excludes parents from number in college, switching it to professional judgment (PJ). (4) Adds examples of other common special circumstances that merit PJ: tuition expenses at an elementary or secondary school, medical or dental expenses not covered by insurance, unusually high child care costs, recent unemployment of a family member, the number of parents enrolled at least half-time in a degree, certificate, or other program leading to a recognized educational credential at a Title IV school, or other changes in the family's income, assets, or student's status. (5) Allows financial aid administrators the authority to refuse to certify a student's loan application on a case by case basis, so long as the school is not discriminating based on race, national origin, religion, sex, marital status, age, or disability status. This allows schools to limit the borrowing of students in specific majors or years in school. It also allows them to refuse to certify a loan if they feel that the student has no intention of repaying the loan. (6) Authorizes the establishment a loan cancellation program for teachers. (7) Authorize the US Department of Education to verify income data submitted on the FAFSA with the IRS. This measure is intended to eliminate a major source of fraud [1, 2].

We collect data from CPS Historical Time Series Tables on School Enrollment (https://www.census.gov/hhes/school/data/cps/historical/index.html). The goal is to compare enrollment rate for different sex of the four age-groups, i.e., 18-19, 20-21, 22-24, and 25-29, before and after the changes of ACT in 1998. To this end, consider the data points (from 1959 to 1998) as the first sample and the last 17 data points (from 1999 to 2015) as the second sample. We wish to answer the following questions:

1. Compute the sample mean and sample covariance matrix of each sample.
2. Are the two samples have the same covariance matrix?
3. Are the mean returns of the two samples equal?

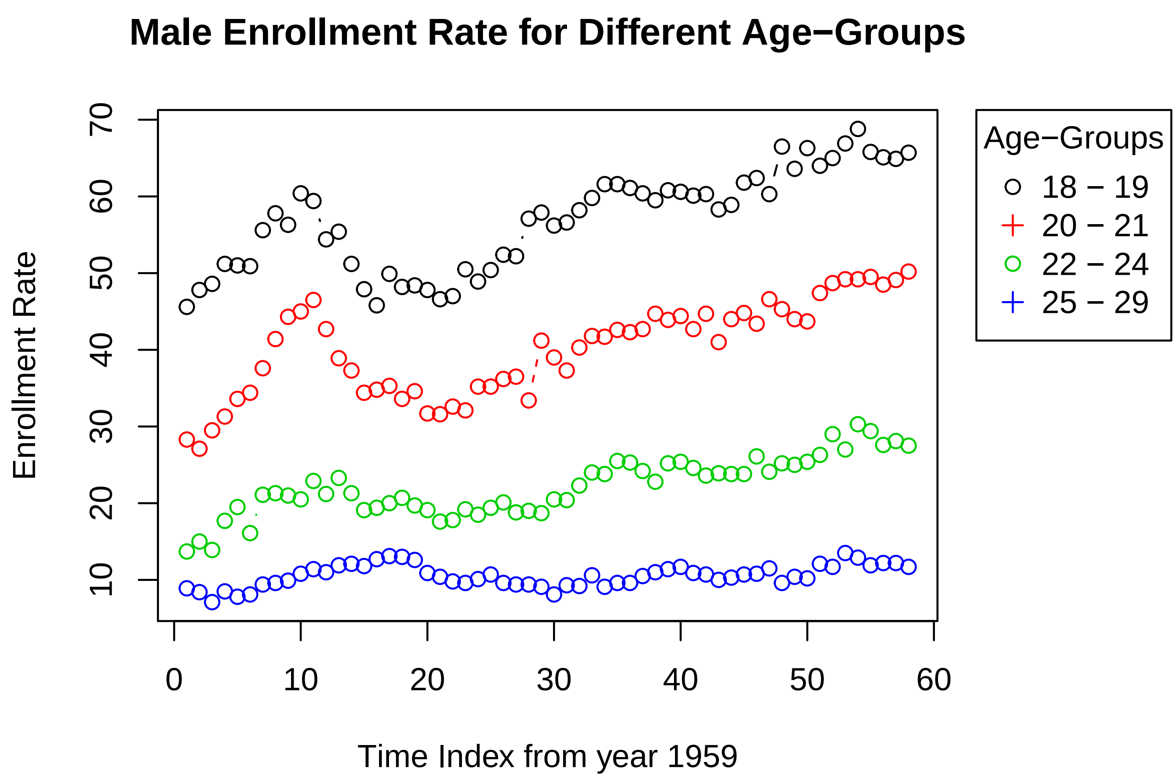
Data Visualization

```r
# Get Data
dataDir="C:\\Data Science App\\W271\\Shih Yu\\Lab3\\Percentage_Population_Univ_Colle_Enrolled_Age_Sex.cs
df <- read.csv(dataDir, header=TRUE)
head(df) # df[,c("A","B","E")]
```

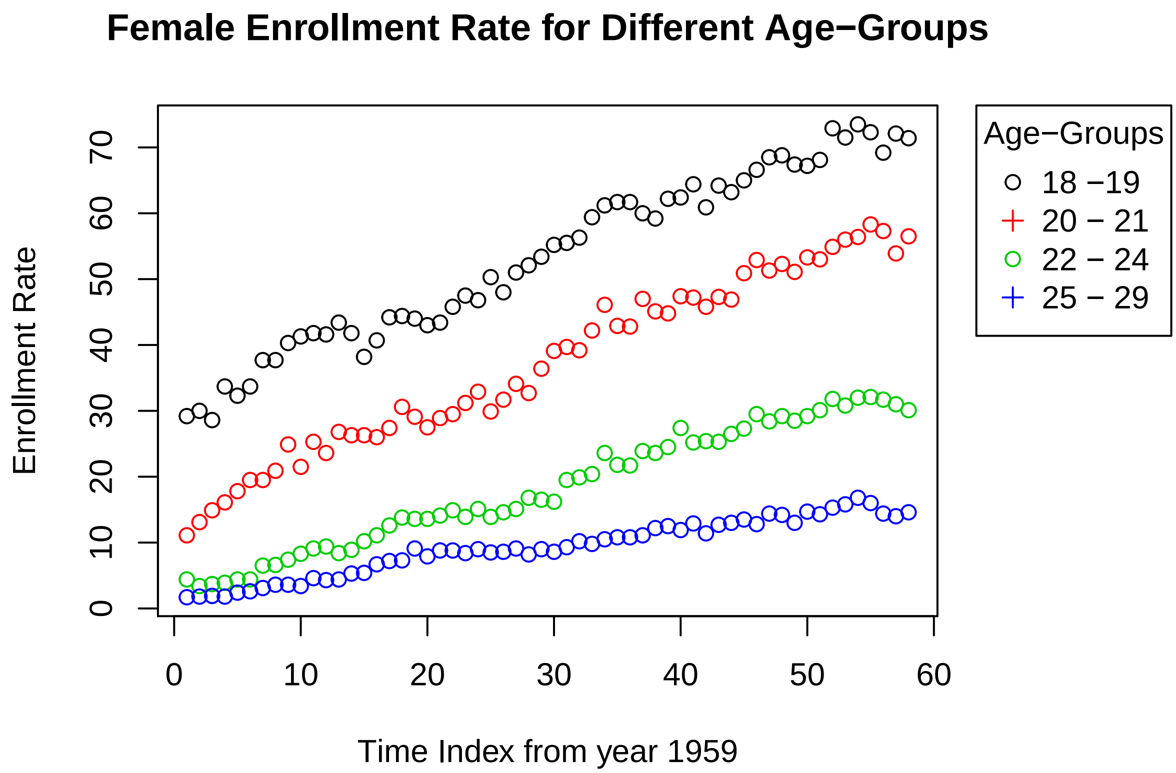
```
##   ï..18_19_years_m X20_21_years_m X22_24_years_m X25_29_years_m
## 1             65.7           50.2           27.5           11.7
## 2             64.9           49.1           28.1           12.2
## 3             65.1           48.5           27.6           12.2
## 4             65.8           49.5           29.4           11.9
## 5             68.8           49.2           30.3           12.9
## 6             66.9           49.2           27.0           13.5
##   X18_19_years_f X20_21_years_f X22_24_years_f X25_29_years_f
## 1           71.4           56.5           30.1           14.6
## 2           72.1           53.9           31.0           14.0
## 3           69.2           57.3           31.7           14.4
## 4           72.3           58.3           32.1           16.0
## 5           73.5           56.4           32.0           16.8
## 6           71.5           56.0           30.8           15.8
```

```r
#str(df)
```

```r
df_m = subset(df, select=c("ï..18_19_years_m", "X20_21_years_m", "X22_24_years_m", "X25_29_years_m"))
df_m = df_m[rev(rownames(df_m)), ]
par(mar=c(5.1, 4.1, 4.1, 8.1), xpd=TRUE)
matplot(df_m, type = c("b"), pch=1,col = 1:4, xlab = "Time Index from year 1959",   ylab = "Enrollment P
legend("topright", inset=c(-0.3, 0), legend = expression("18 - 19", "20 - 21", "22 - 24", "25 - 29"), co
```




```
df_f = subset(df, select=c("X18_19_years_f", "X20_21_years_f", "X22_24_years_f", "X25_29_years_f"))
df_f = df_f[rev(rownames(df_f)), ]
par(mar=c(5.1, 4.1, 4.1, 8.1), xpd=TRUE)
matplot(df_f, type = c("b"), pch=1,col = 1:4, xlab = "Time Index from year 1959",   ylab = "Enrollment R
legend("topright", inset=c(-0.3, 0), legend = expression("18 -19", "20 - 21", "22 - 24", "25 - 29"), col
```



Compute the sample mean and sample covariance matrix of each sample.

```
# Data for male enrollment from 1959 to 1998
x1_m = as.matrix(df[19:58, 1:4])


# Data for female enrollment from 1959 to 1998
x1_f = as.matrix(df[19:58, 5:8])

# Data for male enrollment from 1999 to 2015
x2_m = as.matrix(df[2:18, 1:4])

# Data for female enrollment from 1999 to 2015
x2_f = as.matrix(df[2:18, 5:8])

# mean of male enrollment rate from 1959 to 1998
apply(x1_m,2,mean)
```

```
## ï..18_19_years_m   X20_21_years_m   X22_24_years_m   X25_29_years_m
##           53.825           37.425           20.375           10.180
```

```
# mean of female enrollment rate from 1959 to 1998
apply(x1_f,2,mean)
```

```
## X18_19_years_f X20_21_years_f X22_24_years_f X25_29_years_f
##         46.5175         30.2950         13.5275          7.1050
```

```r
# mean of male enrollment rate from 1999 to 2015
apply(x2_m,2,mean)
```

```
## ï..18_19_years_m   X20_21_years_m   X22_24_years_m   X25_29_years_m
##         63.47059         45.98824         26.07059         11.27059
```

```r
# mean of female enrollment rate from 1999 to 2015
apply(x2_f,2,mean)
```

```
## X18_19_years_f X20_21_years_f X22_24_years_f X25_29_years_f
##       67.98824       52.28235       29.05882       14.07059
```

```r
# covariance matrix of male enrollment rate from 1959 to 1998
cov(x1_m)
```

```
##                  ï..18_19_years_m X20_21_years_m X22_24_years_m
## ï..18_19_years_m       27.3301282      23.293974      12.272692
## X20_21_years_m         23.2939744      26.181923      12.541923
## X22_24_years_m         12.2726923      12.541923       8.726026
## X25_29_years_m         -0.4987179       1.882821       1.722051
##                  X25_29_years_m
## ï..18_19_years_m     -0.4987179
## X20_21_years_m        1.8828205
## X22_24_years_m        1.7220513
## X25_29_years_m        2.1877949
```

```r
# covariance matrix of female enrollment rate from 1959 to 1998
cov(x1_f)
```

```
##                X18_19_years_f X20_21_years_f X22_24_years_f X25_29_years_f
## X18_19_years_f      100.44610       97.88445       64.46284       30.64068
## X20_21_years_f       97.88445       98.96767       64.63835       30.90362
## X22_24_years_f       64.46284       64.63835       44.19538       21.13729
## X25_29_years_f       30.64068       30.90362       21.13729       10.83382
```

```r
# covariance matrix of male enrollment rate from 1999 to 2015
cov(x2_m)
```

```
##                  ï..18_19_years_m X20_21_years_m X22_24_years_m
## ï..18_19_years_m         9.570956       5.862132       5.200331
## X20_21_years_m           5.862132       7.359853       4.732757
## X22_24_years_m           5.200331       4.732757       4.653456
## X25_29_years_m           1.770331       2.484632       1.666581
##                  X25_29_years_m
## ï..18_19_years_m       1.770331
## X20_21_years_m         2.484632
## X22_24_years_m         1.666581
## X25_29_years_m         1.183456
```

```r
# covariance matrix of female enrollment rate from 1999 to 2015
cov(x2_f)
```

```
##                X18_19_years_f X20_21_years_f X22_24_years_f X25_29_years_f
## X18_19_years_f      13.559853      12.429779       8.194485       4.558382
## X20_21_years_f      12.429779      14.342794       8.824228       4.477574
## X22_24_years_f       8.194485       8.824228       5.831324       2.772463
## X25_29_years_f       4.558382       4.477574       2.772463       1.935956
```

From mean value pespective, the male enrollment becomes lower then female at second period (from 1999 to 2015) for all age-groups. For both Sex, students enrollement rates also increase from the first period (from 1959 to 1998) to the second period (from 1999 to 2015) for all age-groups.

Are the two samples have the same covariance matrix?

```r
source("ama.R")
args(BoxM)
```

```
## function (x, nv)
## NULL
```

```r
x_m <- rbind(x1_m,x2_m)
BoxM(x_m, c(40,17))
```

```
## [1] "determinant"
## [1] 341.4437
## [1] "determinant"
## [1] 13.83791
## Test result:
##                 [,1]
## Box.M-C 25.281277183
## p.value  0.004837094
```

Based on the Box-M test statistic, the null hypothesis of equal covariance matrices is rejected. The test statistic is 25.28 with p-value close to zero. For male Sex, the covariance matrix is different for the first sample period (from 1959 to 1998) and the second sample period (from 1999 to 2015).

```r
source("ama.R")
args(BoxM)
```

```
## function (x, nv)
## NULL
```

```r
x_f <- rbind(x1_f,x2_f)
BoxM(x_f, c(40,17))
```

```
## [1] "determinant"
## [1] 488.6356
## [1] "determinant"
## [1] 3.869148
## Test result:
##                  [,1]
## Box.M-C 3.739705e+01
## p.value 4.830532e-05
```

Based on the Box-M test statistic, the null hypothesis of equal covariance matrices is rejected. The test statistic is 37.4 with p-value close to zero. For female Sex, the covariance matrix is different for the first sample period (from 1959 to 1998) and the second sample period (from 1999 to 2015).

Are the mean enrollment of the two samples equal?

```r
args(cmeans)
```

```
## function (da, size, eqV = T, alpha = 0.05)
## NULL
```

```r
cmeans(x_m, c(40,17),eqV=F)
```

```
## [1] "Population 1:"
## Mean-vector:
## ï..18_19_years_m   X20_21_years_m   X22_24_years_m   X25_29_years_m
##             53.8             37.4             20.4             10.2
## Covariance matrix
##                  ï..18_19_years_m X20_21_years_m X22_24_years_m
## ï..18_19_years_m           27.330          23.29          12.27
## X20_21_years_m             23.294          26.18          12.54
## X22_24_years_m             12.273          12.54           8.73
## X25_29_years_m             -0.499           1.88           1.72
##                  X25_29_years_m
## ï..18_19_years_m         -0.499
## X20_21_years_m            1.883
## X22_24_years_m            1.722
## X25_29_years_m            2.188
## [1] "Population 2:"
## Mean-vector:
## ï..18_19_years_m   X20_21_years_m   X22_24_years_m   X25_29_years_m
##             63.5             46.0             26.1             11.3
## Covariance matrix:
##                  ï..18_19_years_m X20_21_years_m X22_24_years_m
## ï..18_19_years_m             9.57           5.86           5.20
## X20_21_years_m               5.86           7.36           4.73
## X22_24_years_m               5.20           4.73           4.65
## X25_29_years_m               1.77           2.48           1.67
##                  X25_29_years_m
## ï..18_19_years_m           1.77
## X20_21_years_m             2.48
## X22_24_years_m             1.67
```

```
## X25_29_years_m             1.18
## differnces in means:
##       [,1]
## [1,] -9.65
## [2,] -8.56
## [3,] -5.70
## [4,] -1.09
## [1] "Hotelling T2, approx-F, & its p-value"
## [1] 8.199977e+01 1.938176e+01 8.200260e-10
## [1] "Simultaneous Tsq. C.I. for difference in means"
## [1] -13.31  -5.98
## [1] -11.99  -5.14
## [1] -8.00 -3.39
## [1] -2.2486  0.0674
## [1] "Simultaneous Bonferroni C.I. for difference in means"
## [1] -12.53  -6.76
## [1] -11.26  -5.86
## [1] -7.51 -3.88
## [1] -2.003 -0.178
## [1] "Critical linear combination: "
##                   [,1]
## ï..18_19_years_m -3.38
## X20_21_years_m   -3.09
## X22_24_years_m   -4.78
## X25_29_years_m    3.96
```

Based on Hotelling T2 test (with un-equal covariances) for male Sex, one can reject the null hypothesis of equal mean vectors at the 5% level. The test statistic is 82.0 with p-value 8.2e-10.

```
args(cmeans)
```

```
## function (da, size, eqV = T, alpha = 0.05)
## NULL
```

```
cmeans(x_f, c(40,17),eqV=F)
```

```
## [1] "Population 1:"
## Mean-vector:
## X18_19_years_f X20_21_years_f X22_24_years_f X25_29_years_f
##          46.52          30.30          13.53           7.11
## Covariance matrix
##                X18_19_years_f X20_21_years_f X22_24_years_f X25_29_years_f
## X18_19_years_f          100.4           97.9           64.5           30.6
## X20_21_years_f           97.9           99.0           64.6           30.9
## X22_24_years_f           64.5           64.6           44.2           21.1
## X25_29_years_f           30.6           30.9           21.1           10.8
## [1] "Population 2:"
## Mean-vector:
## X18_19_years_f X20_21_years_f X22_24_years_f X25_29_years_f
##           68.0           52.3           29.1           14.1
## Covariance matrix:
##                X18_19_years_f X20_21_years_f X22_24_years_f X25_29_years_f
```

```
## X18_19_years_f          13.56          12.43          8.19          4.56
## X20_21_years_f          12.43          14.34          8.82          4.48
## X22_24_years_f           8.19           8.82          5.83          2.77
## X25_29_years_f           4.56           4.48          2.77          1.94
## differnces in means:
##        [,1]
## [1,] -21.47
## [2,] -21.99
## [3,] -15.53
## [4,]  -6.97
## [1] "Hotelling T2, approx-F, & its p-value"
## [1] 1.745064e+02 4.124698e+01 1.554312e-15
## [1] "Simultaneous Tsq. C.I. for difference in means"
## [1] -27.4 -15.5
## [1] -28 -16
## [1] -19.5 -11.6
## [1] -9.00 -4.93
## [1] "Simultaneous Bonferroni C.I. for difference in means"
## [1] -26.2 -16.8
## [1] -26.7 -17.3
## [1] -18.6 -12.4
## [1] -8.57 -5.36
## [1] "Critical linear combination: "
##                  [,1]
## X18_19_years_f   1.45
## X20_21_years_f   4.02
## X22_24_years_f -21.71
## X25_29_years_f   6.19
```

Based on Hotelling T2 test (with un-equal covariances) for feale Sex, one can reject the null hypothesis of equal mean vectors at the 5% level. The test statistic is 174.5 with p-value 1.554312e-15. Comparing male and female, since femal group has larger Hotelling T2 test statistics with smaller p-value, this demonstrates that 1998 Finanicial Aid Act has Sex effect for university or college enrollment rate.

Part II: Time Series Analysis for future University/College Enrollment Rate

```r
# df_m : data frame for male
# df_f : data frame for female
library(gplots)
```

```
## Warning: package 'gplots' was built under R version 3.3.2
```

```
##
## Attaching package: 'gplots'
```

```
## The following object is masked from 'package:stats':
##
##     lowess
```

```r
library(MTS)
```

```
## Warning: package 'MTS' was built under R version 3.3.2
```

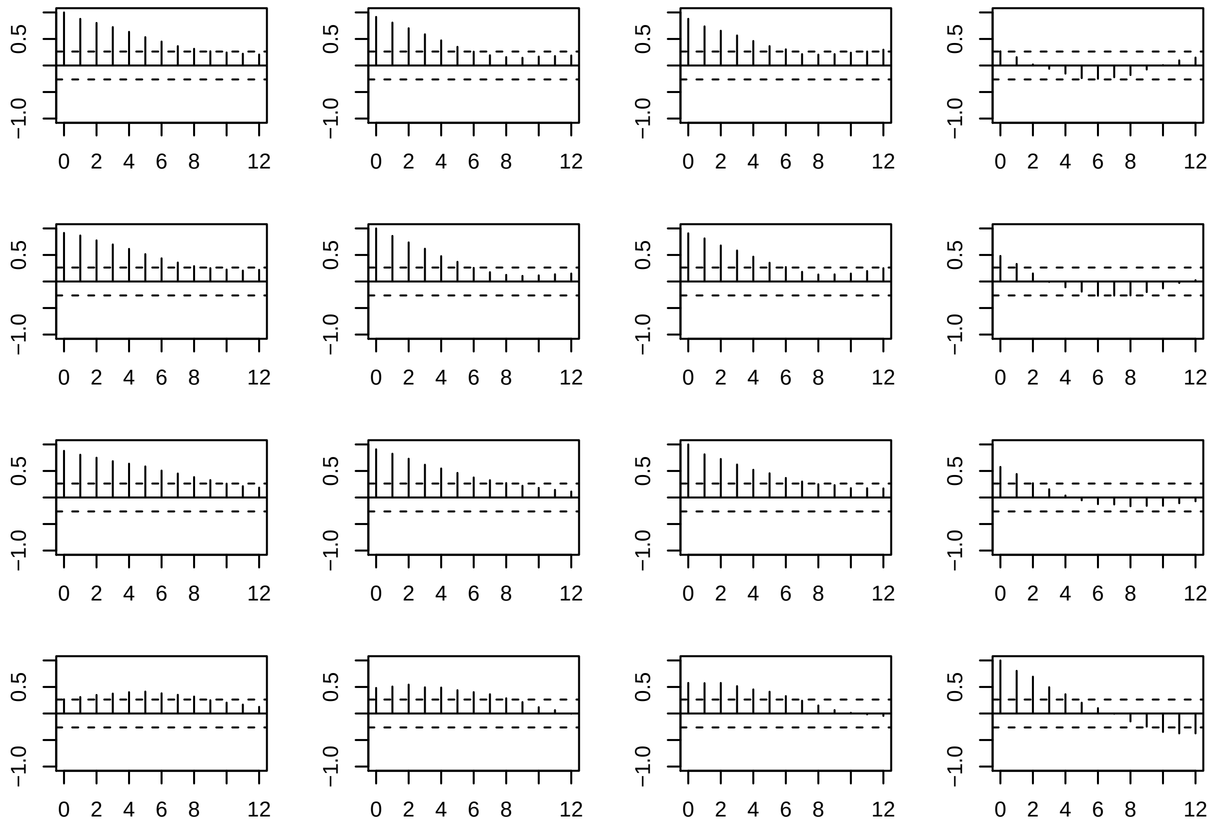
```r
par(mar = rep(2, 4))
dim(df_m)
```

```
## [1] 58  4
```

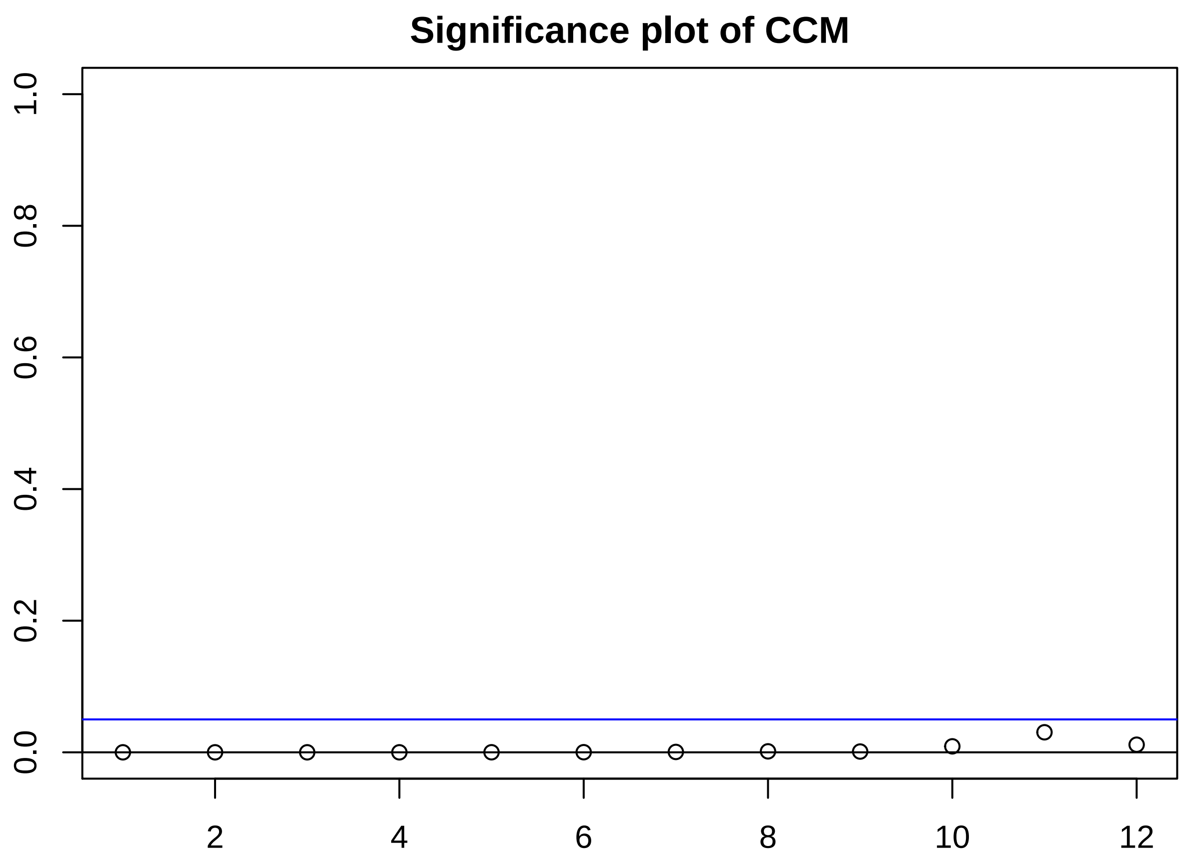
```r
# For numerical stability
rtn = log(df_m[, 1:4] + 2)
#rtn = df_m[, 1:4]
ccm(rtn)
```

```
## [1] "Covariance matrix:"
##                  ï..18_19_years_m X20_21_years_m X22_24_years_m
## ï..18_19_years_m          0.01259        0.01537         0.0161
## X20_21_years_m            0.01537        0.02236         0.0221
## X22_24_years_m            0.01607        0.02211         0.0265
## X25_29_years_m            0.00354        0.00848         0.0110
##                  X25_29_years_m
## ï..18_19_years_m        0.00354
## X20_21_years_m          0.00848
## X22_24_years_m          0.01104
## X25_29_years_m          0.01385
## CCM at lag:  0
##       [,1]  [,2]  [,3]  [,4]
## [1,] 1.000 0.916 0.879 0.268
## [2,] 0.916 1.000 0.908 0.482
## [3,] 0.879 0.908 1.000 0.576
## [4,] 0.268 0.482 0.576 1.000
## Simplified matrix:
## CCM at lag:  1
## + + + .
## + + + +
## + + + +
## + + + +
## CCM at lag:  2
## + + + .
## + + + .
## + + + +
## + + + +
## CCM at lag:  3
## + + + .
## + + + .
## + + + .
## + + + +
## CCM at lag:  4
## + + + .
## + + + .
## + + + .
## + + + +
```

```
## CCM at lag:  5
## + + + .
## + + + .
## + + + .
## + + + .
## CCM at lag:  6
## + . + .
## + . + -
## + + + .
## + + + .
## CCM at lag:  7
## + . . .
## + . . -
## + + + .
## + + . .
## CCM at lag:  8
## + . . .
## + . . .
## + + . .
## + + . .
## CCM at lag:  9
## + . . .
## . . . .
## + . . .
## . . . .
## CCM at lag:  10
## . . . .
## . . . .
## . . . .
## . . . -
## CCM at lag:  11
## . . + .
## . . . .
## . . . .
## . . . -
## CCM at lag:  12
## . . + .
## . . . .
## . . . .
## . . . -
```




```
## Hit Enter for p-value plot of individual ccm:
```




```r
graphics.off()
VARorder(rtn, maxp = 5, output = T)
```

```
## selected order: aic =  1
## selected order: bic =  1
## selected order: hq =  1
## Summary table:
##       p      AIC      BIC       HQ     M(p) p-value
## [1,] 0 -21.5155 -21.5155 -21.5155   0.0000  0.0000
## [2,] 1 -24.8520 -24.2836 -24.6306 184.6883  0.0000
## [3,] 2 -24.8238 -23.6870 -24.3810  22.7750  0.1199
## [4,] 3 -24.7910 -23.0858 -24.1268  20.4982  0.1986
## [5,] 4 -24.5554 -22.2818 -23.6698  11.2204  0.7957
## [6,] 5 -24.5209 -21.6789 -23.4139  16.2929  0.4327
```

The VARorder provides informative order selection. The cross-correlation ma- trices, only the other hand, show that lag-1, to lag-2 have some significant cross- correlations. Also, the simple one-sample t-test shows that the means of the four series are significantly different from zero. Consequently, I start with a VAR(2) model with constant term. The model, however, only employs lag-1 and log-2 coefficient matrix.

```r
apply(rtn,2,t.test)
```

```
## $ï..18_19_years_m
##
```

```
##  One Sample t-test
##
## data:  newX[, i]
## t = 276.15, df = 57, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  4.039484 4.098495
## sample estimates:
## mean of x
##  4.068989
##
##
## $X20_21_years_m
##
##  One Sample t-test
##
## data:  newX[, i]
## t = 189.99, df = 57, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  3.691309 3.769949
## sample estimates:
## mean of x
##  3.730629
##
##
## $X22_24_years_m
##
##  One Sample t-test
##
## data:  newX[, i]
## t = 148.34, df = 57, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  3.129447 3.215094
## sample estimates:
## mean of x
##   3.17227
##
##
## $X25_29_years_m
##
##  One Sample t-test
##
## data:  newX[, i]
## t = 163.13, df = 57, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  2.490126 2.552021
## sample estimates:
## mean of x
##  2.521074
```

```r
# Iteration

n1=VARs(rtn, lags = c(1, 2))#, include.mean = T, fixed = diag(4), prelim = F, details = F, thres = 2)
```

```
## Constant term:
## [1]  0.96655 -0.01664 -0.74541  1.51474
## Std error:
## [1] 0.3960 0.4650 0.6048 0.6778
## AR coefficient matrix:
## AR( 1 )-matrix:
##        [,1]  [,2]    [,3]     [,4]
## [1,]  0.756 0.188 -0.0718 -0.00338
## [2,]  0.688 0.391  0.0264  0.13684
## [3,]  0.622 0.254  0.0980  0.30193
## [4,] -0.374 0.178  0.0736  0.43320
## Standard error:
##       [,1]  [,2]   [,3]   [,4]
## [1,] 0.149 0.113 0.0952 0.0817
## [2,] 0.174 0.133 0.1118 0.0959
## [3,] 0.227 0.173 0.1454 0.1247
## [4,] 0.254 0.193 0.1630 0.1397
##
## AR( 2 )-matrix:
##         [,1]    [,2]  [,3]    [,4]
## [1,] -0.0611 -0.0838 0.197 -0.1958
## [2,] -0.2300  0.0749 0.112 -0.2490
## [3,] -0.1671 -0.0927 0.340 -0.2677
## [4,] -0.3769  0.4039 0.136  0.0572
## Standard error:
##       [,1]  [,2]   [,3]   [,4]
## [1,] 0.173 0.104 0.0867 0.0763
## [2,] 0.203 0.122 0.1018 0.0896
## [3,] 0.264 0.159 0.1324 0.1165
## [4,] 0.296 0.178 0.1484 0.1306
##
## Residuals cov-mtx:
##               [,1]         [,2]         [,3]          [,4]
## [1,]  0.0010445219 0.0003401376 0.0005971094 -0.0001018166
## [2,]  0.0003401376 0.0014400622 0.0001481769  0.0007206703
## [3,]  0.0005971094 0.0001481769 0.0024360135  0.0008987139
## [4,] -0.0001018166 0.0007206703 0.0008987139  0.0030598638
##
## det(SSE) =  6.279228e-12
## AIC =  -24.69033
## BIC =  -23.55353
```

```r
#n1a=refVMAs(n1,thres=0.6)
```

```r
n1a=refVARs(n1,thres=0.5)
```

```
## Constant term:
## [1]  0.8907  0.0000 -0.7454  1.4662
```

```
## Std error:
## [1] 0.2967 0.0000 0.6048 0.5370
## AR coefficient matrix:
## AR( 1 )-matrix:
##        [,1]  [,2]    [,3]  [,4]
## [1,]  0.736 0.178 -0.0733 0.000
## [2,]  0.699 0.397  0.0000 0.139
## [3,]  0.622 0.254  0.0980 0.302
## [4,] -0.324 0.156  0.0000 0.491
## Standard error:
##       [,1]  [,2]   [,3]   [,4]
## [1,] 0.127 0.104 0.0893 0.0000
## [2,] 0.143 0.125 0.0000 0.0847
## [3,] 0.227 0.173 0.1454 0.1247
## [4,] 0.222 0.191 0.0000 0.1072
##
## AR( 2 )-matrix:
##        [,1]    [,2]  [,3]   [,4]
## [1,]  0.000 -0.0928 0.181 -0.188
## [2,] -0.249  0.0847 0.127 -0.254
## [3,] -0.167 -0.0927 0.340 -0.268
## [4,] -0.394  0.4309 0.175  0.000
## Standard error:
##       [,1]  [,2]   [,3]   [,4]
## [1,] 0.000 0.101 0.0756 0.0604
## [2,] 0.157 0.115 0.0817 0.0822
## [3,] 0.264 0.159 0.1324 0.1165
## [4,] 0.290 0.173 0.1353 0.0000
##
## Residuals cov-mtx:
##               [,1]         [,2]         [,3]          [,4]
## [1,]  0.0010468102 0.0003393827 0.0005971094 -0.0001030560
## [2,]  0.0003393827 0.0014422707 0.0001481769  0.0007257571
## [3,]  0.0005971094 0.0001481769 0.0024360135  0.0008987139
## [4,] -0.0001030560 0.0007257571 0.0008987139  0.0030825994
##
## det(SSE) =  6.360043e-12
## AIC =  -24.67754
## BIC =  -23.54074
```

After several iterations with various choices of the threshold in model simplification, the fitted model with AIC about -24.7 due to the smallest AIC value is with constant terms as

$$\begin{bmatrix} 0.9666 \\ -0.0167 \\ -0.7454 \\ 1.51474 \end{bmatrix}$$

and AR(1) coefficients matrix as

$$\begin{bmatrix} 0.756 & 0.188 & -0.0718 & -0.00338 \\ 0.688 & 0.391 & 0.0264 & 0.13684 \\ 0.622 & 0.254 & 0.0980 & 0.30193 \\ -0.374 & 0.178 & 0.0736 & 0.43320 \end{bmatrix}$$

and AR(2) coefficients matrix as

$$\begin{bmatrix} -0.0611 & -0.0838 & 0.197 & -0.1958 \\ -0.2300 & 0.0749 & 0.112 & -0.2490 \\ -0.1671 & -0.0927 & 0.340 & -0.2677 \\ -0.3769 & 0.4039 & 0.136 & 0.0572 \end{bmatrix}$$

```
result = VARpred(n1, h = 10, orig = 0, Out.level = F)
```

```
## orig  58
## Forecasts at origin:  58
##  ï..18_19_years_m X20_21_years_m X22_24_years_m X25_29_years_m
##             4.206          3.925          3.384          2.642
##             4.194          3.916          3.377          2.650
##             4.182          3.899          3.367          2.648
##             4.169          3.882          3.352          2.648
##             4.158          3.868          3.339          2.646
##             4.146          3.853          3.324          2.641
##             4.136          3.840          3.311          2.636
##             4.127          3.827          3.297          2.631
##             4.118          3.816          3.285          2.625
##             4.111          3.805          3.273          2.619
## Standard Errors of predictions:
##          [,1]    [,2]    [,3]    [,4]
##  [1,] 0.03232 0.03795 0.04936 0.05532
##  [2,] 0.04162 0.04985 0.05957 0.06298
##  [3,] 0.04877 0.05640 0.06537 0.07171
##  [4,] 0.05498 0.06299 0.07050 0.07687
##  [5,] 0.06136 0.06981 0.07607 0.07976
##  [6,] 0.06719 0.07669 0.08149 0.08151
##  [7,] 0.07264 0.08330 0.08714 0.08259
##  [8,] 0.07755 0.08957 0.09272 0.08333
##  [9,] 0.08189 0.09529 0.09812 0.08397
## [10,] 0.08565 0.10039 0.10316 0.08464
## Root mean square errors of predictions:
##          [,1]    [,2]    [,3]    [,4]
##  [1,] 0.03474 0.04079 0.05305 0.05945
##  [2,] 6.63328 8.17943 8.43884 7.61589
##  [3,] 6.43247 6.67345 6.81155 8.67634
##  [4,] 6.42373 7.09493 6.67881 7.00841
##  [5,] 6.88879 7.61490 7.22454 5.37877
##  [6,] 6.92782 8.02914 7.39684 4.25718
##  [7,] 6.98674 8.23103 7.80755 3.36840
##  [8,] 6.86576 8.32643 8.01588 2.79318
##  [9,] 6.65961 8.22689 8.12338 2.61871
## [10,] 6.34734 7.99291 8.05666 2.68870
```

Accoding to the above prediction results for future ten years counted from 2015, the enrollement rate of each age-group decreases in the future. The above results should take exponential back to get original enrollement rates measured by percentage. For example, enrollment rate decreaes from $exp(4.206) = 67.08765$ to $exp(4.111) = 61.00769$ at the age group 18 - 19 years old.

References:

[1] http://www.finaid.org/educators/history.phtml [2] https://www.congress.gov/bill/105th-congress/house-bill/6